

Abstract

To extract words or the like intensively related to contents of text from the same text without necessity of cost required for a excessively large amount of man-power and thereby to generate the information of the text using these extracted words or the like. The text information generating method and apparatus comprises an attribute input section for inputting artificial attribute, a discourse structure attribute generating section for generating discourse structure attribute and paragraph length ratio attribute, a combination attribute generating section for generating combination attribute by freely combining artificial attribute, discourse structure attribute and paragraph length ratio attribute, an importance degree estimating section for respectively estimating importance degree indicating enhancement degree of correlation with contents of text for each attribute, a text input interface, an important paragraph determining section for determining the important paragraphs from one or more paragraphs in the input text on the basis of an importance degree of each attribute, and a text output interface for outputting the information of the input text generated on the basis of determination of the important paragraph determining section.